



# RSC811-3

## مهارات بحثية 3

### Task8

**Prepared by**

Mohammed Omar Bin Makhashen

242001540@st.midocean.edu.km

محمد عمر بن مخاشن

Ph.D. student in IT

2025-Apr-5

## المحتوى

### مجتمع وعينة الدراسة

تطبيق الطالب (الباحث) لمفاهيم مجتمع وعينة البحث، مع شرح الأسباب الكامنة وراء القرارات المنهجية المتخذة في مشروع البحث الخاص به.

- مجتمع العينة
- عينة الدراسة
- تصميم العينة
- اخطاء العينة المحتمل
- خاتمة

## "تأثير الخوارزميات الكمومية الهجينة على دقة تصنيف البيانات الطبية"

### المتغيرات:

- الخوارزميات الكمومية الهجينة (Hybrid Quantum Algorithms): الخوارزميات التي تجمع بين الحسابات الكلاسيكية والكمومية لتحقيق أداء أفضل.
- تصنيف البيانات الطبية (Medical Data Classification): استخدام خوارزميات تعلم الآلة لتصنيف البيانات الطبية (مثل الصور الطبية أو بيانات المرضى) لتشخيص الأمراض أو التنبؤ بنتائج العلاج.

## مجتمع وعينة الدراسة: مجتمع البحث

2

## أولاً: مجتمع البحث

**التعريف:** جميع البيانات الطبية المتاحة والقابلة للاستخدام في تطوير وتقييم خوارزميات تصنيف الأمراض والحالات الصحية المختلفة. يشمل هذا التعريف جميع البيانات التي يمكن أن تساهم في تحسين دقة التشخيص وتوجيه العلاج.

## الأنواع:

**بيانات تصويرية:** صور الأشعة السينية (Radiographs)، صور الرنين المغناطيسي (MRI)، صور الأشعة المقطعية (CT scans)، صور الموجات فوق الصوتية (Ultrasound)، صور طب العيون (Ophthalmology Images)، صور الجلدية (Dermatology Images). هذه الأنواع تمثل مجموعة كبيرة من أساليب التصوير الطبي المستخدمة لتشخيص مجموعة واسعة من الحالات.

**بيانات جينية:** بيانات التسلسل الجيني (Genome Sequencing Data)، بيانات التعبير الجيني (Gene Expression Data)، بيانات التعديلات اللاجينية (Epigenetic Modifications Data). توفر هذه الأنواع معلومات حول الأساس الجيني للأمراض والاستجابة للعلاج.

**بيانات سريرية:** السجلات الطبية الإلكترونية (Electronic Health Records - EHR)، نتائج الفحوصات المخبرية (Laboratory Tests)، البيانات الديموغرافية للمرضى (Demographic Data)، التاريخ المرضي للمرضى (Medical History). تسجل هذه الأنواع معلومات مفصلة حول التاريخ الطبي للمرضى وحالاتهم الحالية.

**بيانات نصية:** التقارير الطبية (Medical Reports)، ملاحظات الأطباء (Physician Notes)، استبيانات المرضى (Patient Questionnaires). تتضمن هذه الأنواع معلومات وصفية حول حالة المريض وتجربته.

## مجتمع وعينة الدراسة: مجتمع البحث (استكمال)

3

أولاً: مجتمع البحث (استكمال)

## مبررات الاختيار:

**الأهمية الطبية:** البيانات الطبية هي أساس البحث والتطوير في مجال الرعاية الصحية، وتستخدم لتحسين التشخيص والعلاج.

**التنوع:** يتيح هذا المجتمع الوصول إلى أنواع مختلفة من البيانات الطبية (صور، جينات، بيانات سريرية، نصوص)، مما يسمح بتقييم أداء الخوارزميات في سياقات متعددة.

**التوفر:** تتوفر العديد من مجموعات البيانات الطبية المتاحة للعامة (Public Datasets)، مما يسهل الوصول إلى البيانات اللازمة لإجراء البحث ويضمن الشفافية وإمكانية تكرار النتائج.

**التمثيلية لمجال الرعاية الصحية ككل:** من خلال تغطية مجموعة واسعة من الأمراض والحالات الصحية وأساليب التشخيص، يمثل هذا المجتمع مجال الرعاية الصحية ككل.

## مجتمع وعينة الدراسة: عينة الدراسة

## 4

## ثانياً: عينة الدراسة

**التعريف:** مجموعة فرعية محددة من مجتمع البحث تم اختيارها لتمثيل المجتمع الأكبر وتستخدم بشكل عملي في هذه الدراسة. تهدف العينة إلى توفير بيانات قابلة للتحليل والتطبيق في تطوير وتقييم الخوارزميات.

## المكونات:

**مجموعات بيانات صور الأشعة السينية لمرض السل (Tuberculosis):** تم اختيار هذه المجموعات لتمثيل تحدي تصنيف الأمراض المعدية، ولتنوع صور الأشعة السينية المتاحة.

**مجموعات بيانات صور سرطان الثدي (Breast Cancer):** تم اختيار هذه المجموعات لتمثيل تحدي تصنيف الأمراض السرطانية، ولأهمية الكشف المبكر عن سرطان الثدي.

**مجموعات بيانات التصوير بالرنين المغناطيسي للدماغ (Brain MRI):** تم اختيار هذه المجموعات لتمثيل تحدي تصنيف الأمراض العصبية، ولتعقيد صور الرنين المغناطيسي للدماغ.

## مجتمع وعينة الدراسة: عينة الدراسة (استكمال)

5

ثانياً: عينة الدراسة (استكمال)

مبررات الاختيار:

**التمثيلية:** تمثل العينة أنواعًا مختلفة من البيانات الطبية (صور الأشعة السينية، صور الرنين المغناطيسي) والأمراض والحالات الصحية المختلفة (مرض السل، سرطان الثدي، تشوهات الدماغ)، مما يزيد من إمكانية تعميم النتائج.

**الجودة:** تتميز مجموعات البيانات المختارة بجودتها العالية وتوفرها على معلومات دقيقة وموثوقة، مما يضمن جودة النتائج.

**التوفر:** مجموعات البيانات المختارة متاحة للعامة، مما يسهل الوصول إليها واستخدامها في البحث، ويسهل عملية التحقق من النتائج.

**القدرة على الإجابة على أسئلة البحث:** تم اختيار مجموعات البيانات التي تسمح بمعالجة أسئلة البحث بشكل مباشر، من خلال توفير البيانات اللازمة لاختبار الفرضيات.

## مجتمع وعينة الدراسة: تصميم العينة

6

## ثالثاً: تصميم العينة

**حجم العينة:** تم تحديد حجم العينة بحيث يتضمن ما لا يقل عن 1000 صورة لكل نوع من أنواع البيانات (صور الأشعة السينية، صور الرنين المغناطيسي، إلخ).

## مبررات حجم العينة:

- **القوة الإحصائية:** يتضمن هذا الحجم الحصول على قوة إحصائية كافية للكشف عن أي اختلافات ذات دلالة إحصائية بين أداء الخوارزميات الكمومية الهجينة والخوارزميات الكلاسيكية.
- **التنوع:** يسمح هذا الحجم بالتقاط التنوع الموجود داخل كل نوع من أنواع البيانات، مما يزيد من تمثيلية العينة.
- **الموارد المتاحة:** تم تحديد هذا الحجم بناءً على الموارد المتاحة لإجراء البحث، مع تحقيق التوازن بين الحصول على نتائج دقيقة واستخدام الموارد بكفاءة.
- **تقليل أخطاء المعاينة المحتملة:** يساهم الحجم الكبير نسبياً في تقليل تأثير أخطاء المعاينة العشوائية على النتائج.

## مجتمع وعينة الدراسة: تصميم العينة (استكمال)

7

## ثالثاً: تصميم العينة (استكمال)

**طريقة أخذ العينات:** تم استخدام العينة العشوائية البسيطة (Simple Random Sampling) من كل مجموعة بيانات متاحة.

## مبررات الاختيار:

- **تقليل التحيز:** تضمن هذه الطريقة أن يكون لكل عنصر (صورة) في المجموعة فرصة متساوية للاختيار، مما يقلل من التحيز ويزيد من تمثيلية العينة.
- **سهولة التنفيذ:** تعتبر هذه الطريقة سهلة التنفيذ نسبياً، مما يجعلها خياراً عملياً في هذا البحث.
- **الأساس النظري:** توفر هذه الطريقة أساساً نظرياً واضحاً لتقييم النتائج الإحصائية.

## مجتمع وعينة الدراسة: أخطاء المعاينة المحتملة

8

## رابعاً: أخطاء المعاينة المحتملة

**التحيز في البيانات:** قد تحتوي مجموعات البيانات الطبية المتاحة على تحيز في البيانات، مثل تمثيل غير متساو للسكان المختلفين (التحيز الديموغرافي) أو وجود بيانات مضللة أو غير دقيقة (التحيز في القياس). يجب مراعاة هذه التحيزات عند تحليل النتائج وتفسيرها.

**القيود التقنية:** قد تحد القيود التقنية (مثل قوة المعالجة المتاحة) من القدرة على معالجة وتحليل كميات كبيرة من البيانات، مما قد يؤثر على حجم العينة أو جودة التحليل.

**أخطاء المعاينة:** على الرغم من استخدام العينة العشوائية البسيطة، لا يمكن تجنب أخطاء المعاينة العشوائية تمامًا. من المحتمل أن العينة المختارة لا تمثل تمامًا جميع خصائص مجتمع البحث، مما قد يؤثر على النتائج. ولتقليل هذا التأثير، تم اختيار حجم عينة كبير نسبيًا واستخدام تقنيات إحصائية مناسبة لتحليل البيانات.

**خاتمة: مجتمع وعينة الدراسة**

يمثل اختيار مجتمع وعينة الدراسة خطوة حاسمة في أي بحث علمي، حيث يحدد هذا الاختيار نطاق النتائج وقابليتها للتعميم. في هذا الفصل، تم توضيح الأسس التي تم الاستناد إليها لتحديد مجتمع البحث وعينة الدراسة في هذا البحث، مع التركيز على أهمية تحقيق التوازن بين التمثيلية والجودة والجدوى العملية.

من خلال استخدام مجموعات البيانات الطبية المتاحة للعامة، يهدف هذا البحث إلى ضمان الشفافية وإمكانية تكرار النتائج، بالإضافة إلى الوصول إلى كميات كبيرة من البيانات الطبية المتنوعة. ومن خلال استخدام العينة العشوائية البسيطة، يهدف البحث إلى التقليل من التحيز وزيادة تمثيلية العينة لمجتمع البحث الأكبر.

على الرغم من وجود بعض القيود والمشاكل المحتملة، إلا أن الباحث يرى أن المنهجية المتبعة في اختيار مجتمع وعينة الدراسة تساهم بشكل كبير في تحقيق أهداف هذا البحث، وتقديم رؤى قيمة حول إمكانات الخوارزميات الكمومية الهجينة في تصنيف البيانات الطبية.



# Thank You